

# SHIVAM LALAKIYA

☎ 781-363-8482 ✉ [lalakiya.s@northeastern.edu](mailto:lalakiya.s@northeastern.edu) 🔗 [linkedin.com/in/shivam-lalakiya](https://www.linkedin.com/in/shivam-lalakiya) 🐙 [github.com/shivamlalakiya](https://github.com/shivamlalakiya)

## Education

**Northeastern University, Boston, MA**

**Sept 2021 – July 2023**

*Master of Science in Data Analytics Engineering*

*GPA: 3.8/4*

**Sardar Vallabhbhai National Institute of Technology (NIT Surat), India**

**July 2017 – May 2021**

*Bachelor of Technology in Electronics and Communication Engineering*

*GPA: 8.1/10*

## Technical Skills

**Languages & Frameworks:** Python, R, SQL, MATLAB, Shell scripting, NoSQL, C/C++, Java

**Tools & Technologies:** Machine Learning, NLP, Computer Vision, Generative-AI, LLMs, Transformer, Power BI, Tableau, ETL, Airflow, AWS, GIT, CI/CD, Kafka, Hadoop, Docker, LaTeX

## Experience

**Invicro, LLC.**

**Sept 2023 – Present**

*Associate Data Manager*

*Boston, MA*

- Developed an **automated data pipeline** with Apache Airflow for transferring pathological images, reducing data processing time by 20% and ensuring reliable, on-time data delivery for research studies.
- Engineered and deployed **Computer Vision Models**, resulting in a 10 % improvement in pathological image analysis and bio-marker Identification, translating into 15 % of more study sponsors.
- Established **data quality** control procedures, validation protocols, and **data auditing** frameworks to improve data precision and ensure conformity with industry benchmarks, resulting in dependable research outcomes and reports.
- Collaborated with lab scientists to automate and standardize slide and file naming, eliminating errors by 17%.

**Ring Therapeutics**

**May 2022 – Dec 2022**

*Data Science Co-op*

*Boston, MA*

- Provided **analytical insights** on binding sites and cell specificity for sequenced data collected by Discovery team from protein sequences, enhancing outcomes and advancing Anello-backed programmable medicine.
- Leveraged positional encoded **NLP-based models** for gene/protein sequences and **Graph Neural Networks** for predicting tropism, achieving a 0.75 F1-score to generate vectors for safe delivery of therapies to target cells.
- Fine-tuned the **Protein-BERT** model with Anello sequences, attaining a notable 0.8 F1 score for accurate tropism classification and motif identification utilizing MEME-based analysis techniques.
- Built **ETL pipeline & Docker container** with **Airflow DAG** to automate 90% of data loading into fasta format.

**IIT Madras**

**Jan 2020 – Jan 2021**

*Research Assistant*

*Chennai, India*

- Executed a 'Caching with Deep Learning' project, leveraging **RNNs** and **LSTM** models for time-series analysis of user preferences, resulting in a 90 % success rate in accurately forecasting future requests.
- Preprocessed and filtered 100s GB data using Pandas, employing advanced **data manipulation** for optimal training.
- Designed an LSTM-based caching policy, outperforming traditional policies (LIFO, LRU, and LFU) with a **130% improvement in hit rates**, as demonstrated through extensive benchmarking and performance analysis.

**IIT Madras**

**Jan 2020 – Jan 2021**

*Research Assistant*

*Chennai, India*

- Developed a **Multimodal Biometric System** using Convolutional Neural Networks (CNNs) to combine Iris, Facial, Speech Recognition, and fingerprints, achieving an 85% precision rate on multiple datasets
- Deployed the model on **cloud platforms** (AWS and GCP) for real-time authentication, scalability and accessibility.
- Conducted testing and validation, achieving an F1 score of 0.87, ensuring the system's reliability and precision.
- Applied the system to **automate attendance** in departmental classes, saving 15 minutes per class session and enhancing security in the professor's cabin.

## Projects

**Analytics and Visualization using R programming** | *R, SQL, Tableau, ETL, A/B testing* | [Code](#) **Sep – Dec 2021**

- Performed clustering, probabilistic evaluation, and A/B testing to find insights on University and E-commerce dataset.
- Derived vital factors influencing top 100 university rankings and reasons for hierarchy shifts.
- Analyzed discounts, sales, and profit correlation for e-commerce websites, revealing festive seasons exhibit a 35% increase in sales, while non-festive seasons demonstrate a 30% rise in deals.

## Extra-Curriculum

- Student Success Guide at Northeastern University:** Advised and mentored 100+ students per semester for meaningful academic experiences and effective stress management, propelling 85% of students towards career advancement.
- Math Tutor:** Tutored calculus, linear algebra, probability, etc., boosting student understanding and achievement.